# RETAKH'S MOTZKIN PATHS AND SOME COMBINATORIAL COMMENTS

HELMUT PRODINGER

ABSTRACT. Dyck paths where peaks are only allowed on level 1 and on even-indexed levels, were introduced by Retakh and analysed by Zeilberger, with assistance from Ekhad. We add some combinatorial comments to the enumeration, which involves Motzkin numbers, in particular, about the average height of such objects.

1. INTRODUCTION

V. Retakh [3] introduced the following restricted class of Dyck paths: Peaks are only allowed on level 1 and on even-numbered levels. Here is an example, and the corresponding plane tree using the standard bijection.
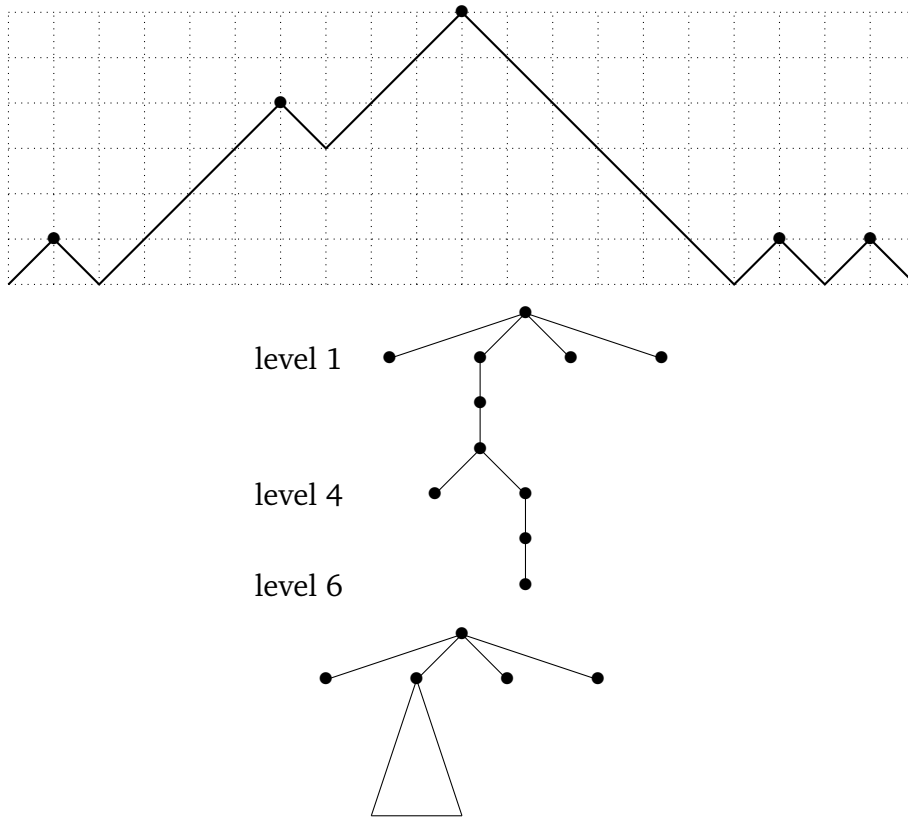
Ekhad and Zeilberger [3] proved recently that these restricted paths are enumerated by Motzkin numbers. Recall that the generating function of the Motzkin numbers $M(z)$ according to length satisfies $M = 1 + zM + z^2M^2$ and thus

$$M(z) = \frac{1 - z - \sqrt{1 - 2z - 3z^2}}{2z^2}.$$

In this note, I want to present a few additional observations, also including the height of the paths (or the associated plane trees). First, we are going to confirm the connection to Motzkin paths:

Since the level 1 is somewhat special, we only consider trees as symbolized by the triangle. We will use two generating functions, to deal with the odd/even situation. We have

$$F = \frac{zG}{1 - G} \quad \text{and} \quad G = \frac{z}{1 - F}.$$

Solving, we find $F(z) = z^2 M(z)$ and the total generating function is

$$\frac{z}{1 - z} \sum_{r \geq 0} \left(\frac{F}{1 - z} \right)^r = zM(z),$$

as predicted. Recall that the number of nodes in trees is always one more than the half-length of the corresponding Dyck path.

We will compute the average height of such restricted paths, using singularity analysis of generating functions, as in [4, 5]. Whether we define the height in terms of the maximal chain of edges resp. nodes only makes a difference of one, and we will only compute the average height according to the leading term of order $\sqrt{n}$. For readers who wish to see how more terms could be computed, at least in principle, we cite [8].

2. THE HEIGHT

Now we will use the substitution $z = \frac{v}{1 + v + v^2}$, which occured for the first time in [9], but has been used more recently in different models where Motzkin numbers are involved [10, 7, 1]. For example, we have simply $M(z) = 1 + v + v^2$. We define

$$G_{k+1} = \frac{z}{1 - \dfrac{zG_k}{1 - G_k}}, \quad \text{with} \quad G_1 = z.$$

There is a simple formula, viz.

$$G_k = \frac{v}{1 + v} \frac{1 - v^{2k}}{1 - v^{2k+1}}.$$

This is easy to prove by induction, which we will do for the convenience of the reader. The start is

$$G_1 = \frac{v}{1 + v} \frac{1 - v^2}{1 - v^3} = \frac{v}{1 + v} \frac{1 + v}{1 + v + v^2} = \frac{v}{1 + v + v^2} = z.$$

And now

$$G_{k+1} = \frac{z}{1 - \dfrac{zG_k}{1-G_k}} = \frac{z(1-G_k)}{1-(1+z)G_k} = \frac{v}{1+v+v^2}\frac{1 - \frac{v}{1+v}\frac{1-v^{2k}}{1-v^{2k+1}}}{1 - \frac{(1+v)^2}{1+v+v^2}\frac{v}{1+v}\frac{1-v^{2k}}{1-v^{2k+1}}}$$

$$= \frac{v}{1+v}\frac{1+v-v\frac{1-v^{2k}}{1-v^{2k+1}}}{1+v+v^2-v(1+v)\frac{1-v^{2k}}{1-v^{2k+1}}} = \frac{v}{1+v}\frac{(1+v)(1-v^{2k+1})-v(1-v^{2k})}{(1+v+v^2)(1-v^{2k+1})-v(1+v)(1-v^{2k})}$$

$$= \frac{v}{1+v}\frac{1-v^{2k+2}}{1-v^{2k+3}},$$

as claimed. From this we also get

$$F_k = \frac{zG_k}{1-G_k} = \frac{v^2}{1+v+v^2}\frac{1-v^{2k}}{1-v^{2k+2}}.$$

For $k \geq 1$, F_k is the generating function of trees (like in the triangle) of height $\leq 2k$.

Note that the height is currently counted in terms of nodes;

$$F_1 = \frac{z^2}{1-z},$$

which describes a root with $\ell \geq 1$ leaves attached to the root.

Now we incorporate the irregular beginning of the tree and compute

$$\frac{z}{1-z}\sum_{r\geq 0}\left(\frac{F_h}{1-z}\right)^r = \frac{z}{1-z}\frac{1}{1-\dfrac{F_h}{1-z}} = v\frac{1-v^{2h+2}}{1-v^{2h+4}}.$$

From here onwards it seems to be more natural to define the height of the whole tree in terms of the number of *edges*, and then the quantity we just derived is the generating function of all trees with height $\leq 2h$, for $h \geq 1$. Note that the limit $h \to \infty$ gives us simply $v = zM(z)$, which is consistent. There is also a contribution of trees of height ≤ 1, namely $\frac{z}{1-z} = \frac{v}{1+v^2}$, but this term is, when we compute the average height, irrelevant and only contributes to the error term, as we only compute the leading term, which is of order $\sqrt{n}$.

So, apart from normalization, we are led to investigate

$$\sum_{h\geq 1}2h\left[v\frac{1-v^{2h+2}}{1-v^{2h+4}} - v\frac{1-v^{2h}}{1-v^{2h+2}}\right] = 2v(1-v^{-2})\sum_{h\geq 1}h\left[\frac{v^{2h+4}}{1-v^{2h+4}} - \frac{v^{2h+2}}{1-v^{2h+2}}\right]$$

$$= 2v(1-v^{-2})\sum_{h\geq 0}h\frac{v^{2h+4}}{1-v^{2h+4}} - 2v(1-v^{-2})\sum_{h\geq 0}(h+1)\frac{v^{2h+4}}{1-v^{2h+4}}$$

$$= -2v + \frac{2(1-v^2)}{v}\sum_{h\geq 1}\frac{v^{2h}}{1-v^{2h}}.$$

Note that we could get explicit coefficients fr0m here, using trinomial coefficient, $\binom{n,3}{k} = [v^k](1+v+v^2)^n$ (notation from [2]). To show the reader how this works, we compute

$$[z^{n+1}]\frac{1-v^2}{v}\sum_{h\geq 1}\frac{v^{2h}}{1-v^{2h}} = \frac{1}{2\pi i}\oint\frac{dz}{z^{n+2}}\frac{1-v^2}{v}\sum_{h\geq 1}\frac{v^{2h}}{1-v^{2h}}$$

$$= \frac{1}{2\pi i}\oint dv(1-v^2)^2\frac{(1+v+v^2)^n}{v^{n+3}}\sum_{h\geq 1}\sum_{h,k\geq 1}v^{2hk}$$

$$= [v^{n+2}](1-2v^2+v^4)\sum_{h\geq 1}d(h)v^{2h}(1+v+v^2)^n$$

$$= \sum_{h\geq 1}d(h)\left[\binom{n,3}{n+2-2h}-2\binom{n,3}{n-2h}+\binom{n,3}{n-2-2h}\right].$$

Note that $d(h)$ is the number of divisors of h. We will, however, not use this explicit form. The expression as derived before,

$$-2v+\frac{2(1-v^2)}{v}\sum_{h\geq 1}\frac{v^{2h}}{1-v^{2h}},$$

has to be expanded around $v=1$, which is a standard application of the Mellin transform. Details are worked out in [6], for example:

$$\sum_{h\geq 1}\frac{v^{2h}}{1-v^{2h}} = \sum_{k\geq 1}d(k)v^{2k} \sim -\frac{\log(1-v^2)}{1-v^2} \sim -\frac{\log(1-v)}{2(1-v)}.$$

Note again that $d(k)$ is the number of divisors of k. Consequently

$$-2v+\frac{2(1-v^2)}{v}\sum_{h\geq 1}\frac{v^{2h}}{1-v^{2h}} \sim -2\log(1-v).$$

We have $1-v \sim \sqrt{3}\sqrt{1-3z}$, and $z = \frac{1}{3}$ is the relevant singularity when discussing Motzkin numbers. We can continue

$$-2v+\frac{2(1-v^2)}{v}\sum_{h\geq 1}\frac{v^{2h}}{1-v^{2h}} \sim -\log(1-3z).$$

The coefficient of z^n in this is $\frac{3^n}{n}$. This has to be divided by

$$[z^n]zM(z) = [z^{n-1}]M(z) \sim \frac{3^{n+\frac{1}{2}}}{2\sqrt{\pi}n^{3/2}},$$

with the final result for the average height of the restricted Dyck paths (à la Retakh):

$$\sim 2\sqrt{\frac{\pi n}{3}}.$$

Recall [9] that the average height of Motzkin paths of length n is asymptotic to

$$\sqrt{\frac{\pi n}{3}}.$$

3. The number of leaves

We can use a second variable, u, to count the number of leaves. Then we have

$$F(z,u) = \frac{zG(z,u)}{1-G(z,u)} \quad \text{and} \quad G(z,u) = zu + \frac{zF(z,u)}{1-F(z,u)},$$

which leads to

$$F(z,u) = \frac{1-zu-z^2+z^2u-\sqrt{1-2zu-2z^2-2z^2u+z^2u^2-2z^3u+2z^3u^2+z^4-2z^4u+z^4u^2}}{2(1-zu+z)}.$$

Bringing the irregular beginning also into the game leads to

$$\frac{z}{1-zu}\sum_{r\geq 0}\left(\frac{F}{1-zu}\right)^r + zu - z.$$

This is an ugly expression that we do not display here. But we can compute the average number of leaves, by differentiation w.r.t. u, followed by setting $u=1$:

$$R := \frac{v(1+v)(1-v+2v^2-v^3)}{(1-v)(1+v+v^2)}.$$

The coefficient of z^n in this can be expressed in terms of trinomial coefficients, if needed. But we only compute an asymptotic formula, to keep this section short. Expanding around $v=1$, we find

$$R \sim \frac{2}{3}\frac{1}{1-v} \sim \frac{2}{3}\frac{1}{\sqrt{3}\sqrt{1-3z}},$$

and thus

$$[z^n]R \sim \frac{2}{3}\frac{1}{\sqrt{3}}3^n\frac{1}{\sqrt{\pi n}}.$$

We divide this again by

$$\frac{3^{n+\frac{1}{2}}}{2\sqrt{\pi}n^{3/2}}$$

with the result

$$\frac{4}{9}n,$$

which is the asymptotic number of leaves in a Retakh tree of size n. Recall that for unrestricted planar trees, the result is $\frac{n}{2}$, which is a folklore result using Narayana numbers. So the constant in the restricted case, $\frac{4}{9}$, is a bit smaller than $\frac{1}{2}$.

With some effort, more precise approximations could be obtained, as well as the variance. This might be a good project for a student.

4. Conclusion

I am always happy to see a new occurrence of Motzkin numbers and that the methods that I learnt more than 40 years ago from Knuth, Flajolet, and others still work.

References

[1] Benjamin Hackl, Clemens Heuberger, and Helmut Prodinger. The *B*-project. 2020.
[2] Louis Comtet. *Advanced combinatorics*. D. Reidel Publishing Co., Dordrecht, enlarged edition, 1974. The art of finite and infinite expansions.
[3] Shalosh B. Ekhad, Doron Zeilberger, Automatic Counting of Restricted Dyck Paths via (Numeric and Symbolic) Dynamic Programming arXiv:2006.01961, 2020
[4] Philippe Flajolet and Andrew Odlyzko. Singularity analysis of generating functions. *SIAM J. Discrete Math.*, 3(2):216–240, 1990.
[5] P. Flajolet and R. Sedgewick, Analytic Combinatorics, Cambridge University Press, 2009.
[6] Clemens Heuberger, Helmut Prodinger, and Stephan Wagner. The height of multiple edge plane trees. *Aequationes Math.*, 90(3):625–645, 2016.
[7] Benjamin Hackl, Clemens Heuberger, Helmut Prodinger, Ascents in Non-Negative Lattice Paths, arXiv:1801.02996, 2018.
[8] Helmut Prodinger. The height of planted plane trees revisited. *Ars Combinatoria*, 16-B:51-55, 1983.
[9] Helmut Prodinger. The average height of a stack where three operations are allowed and some related problems. *J. Combin. Inform. System Sci.*, 5(4):287–304, 1980.
[10] H. Prodinger, Deutsch paths and their enumeration, submitted (2020).

Helmut Prodinger, Department of Mathematical Sciences, Stellenbosch University, 7602 Stellenbosch, South Africa, and, Department of Mathematics and Mathematical Statistics, Umea University, 907 36 Umea, Sweden

Email address: `hproding@sun.ac.za`